The structure of the symptoms of major depression: exploratory and confirmatory factor analysis in depressed Han Chinese women


1Wellcome Trust Centre for Human Genetics, Oxford, UK; 2Virginia Institute for Psychiatric and Behavioral Genetics, Department of Psychiatry, Virginia Commonwealth University, Richmond, VA, USA; 3Shanghai Mental Health Center, Shanghai, P.R. China (PRC); 4Huashan Hospital of Fudan University, Shanghai, PRC; 5Chinese Traditional Hospital of Zhejiang, Hangzhou, Zhejiang, PRC; 6No. 1 Hospital of Zhejiang University, Hangzhou, Hangzhou, PRC; 7Xinhua Hospital of Zhejiang Province, Hangzhou, Zhejiang, PRC; 8No. 1 Hospital of Shanxi Medical University, Taiyuan, Shanxi, PRC; 9Shenyang Hospital of China Medical University, Shenyang District, Shenyang, Liaoning, PRC; 10No. 1 Hospital of Medical College of Xian Jiaotong University, Xian, Shaanxi, PRC; 11Jilin Brain Hospital, Siping, Jilin, PRC; 12The First Hospital of China Medical University, Shenyang, Liaoning, PRC; 13Mental Hospital of Jiangxi Province, Nanchang, Jiangxi, PRC; 14Xiamen Mental Health Center, New Qujiang District, Xian, Shaanxi, PRC; 15Beijing Anding Hospital of Capital University of Medical Sciences, Deshengmen wai, Xicheng District, Beijing, PRC; 16Shandong Mental Health Center, Jinan, Shandong, PRC; 17No. 3 Hospital of Sun Yat-sen University, Tianhe District, Guangzhou, Guangdong, PRC; 18Hebei Mental Health Center, Baoding, Hebei, PRC; 19Chongqing Mental Health Center, Jiangbei District, Chongqing, PRC; 20Tianjin Anding Hospital, Hexi District, Tianjin, PRC; 21No. 4 Hospital of Jangsu University, Zhenjiang, Jiangsu, PRC; 22Psychiatric Hospital of Hubei Province, Xinci, Hanen, PRC; 23Nanjing Brain Hospital, Nanjing, Jiangsu, PRC; 24Harbin Medical University, Nangang District, Harbin, Heilongjiang, PRC; 25Shenzhen Kang Ning Hospital, Luolu District, Shenzhen, Guangdong, PRC; 26First Hospital of Hebei Medical University, Shijiazhuang, Hebei, PRC; 27Guangzhou Brain Hospital (Guangzhou Psychiatric Hospital), Liwan District, Guangzhou, Guangdong, PRC; 28No. 1 Hospital of Chongqing Medical University, Yuzhong District, Chongqing, PRC; 29Dahua No. 7 Hospital, Ganjiaying District, Dalian, Liaoning, PRC; 30No. 3 Hospital of Heilongjiang Province, Beian, Heilongjiang, PRC; 31Wuhan Mental Health Center, Wuhan, Hubei, PRC; 32Sichuan Mental Health Center, Mianyang, Sichuan, PRC; 33Mental Health Institute of Jining Medical College, Dai Zhaung, Bei Jiao, Jining, Shandong, PRC; 34Liucheng No. 4 Hospital, Liucheng, Shandong, PRC; 35Mental Health Center of West China Hospital of Sichuan Medical University, Wuwu District, Chengdu, Sichuan, PRC; 36Suzhou Guangji Hospital, Suzhou, Jiangsu, PRC; 37Anhui Mental Health Center, Hefei, Anhui, PRC; 38Xinbo Kang Ning Hospital, Zhenhai District, Ningbo, Zhejiang, PRC; 39Changzhou Mental Hospital, Changzhou, Jilin, PRC; 40No. 2 Hospital of Lanzhou University, Lanzhou, Gansu, PRC; 41Fuzhou Psychiatric Hospital, Cangzhou District, Fuzhou, Fujian, PRC; 42Harbin No. 1 Special Hospital, Harbin, Heilongjiang, PRC; 43Jining Psychiatric Hospital, North Dai Zhaung, Rencheng District, Jining, Shandong, PRC; 44No. 2 Xiangya Hospital of Zhongnan University, Furong District, Changsha, Hunan, PRC; 45XiJing Hospital of No. 4 Military Medical University, Xian, Shaanxi, PRC; 46Mental Health Center of Shantou University, Shantou, Guangdong, PRC; 47Tongji University Hospital, Shanghai, PRC; 48Huairou No. 3 Hospital, Huairou, Beijing, PRC; 49Wuhan No. 3 Hospital, Huayang, Hubei, PRC; 50Mudanjiang Psychiatric Hospital of Heilongjiang Province, Xinglong, Mudanjiang, Heilongjiang, PRC; 51No. 1 Hospital of Jinan University, Guangzhou, Guangdong, PRC; 52Qingtou Mental Health Center, Shibi District, QinDa, Shandong, PRC; 53Guangxi Longquan Hospital, Yufeng District, Liuzhou, PRC; 54Daping No. 3 Hospital of Heilongjiang Province, Rongjiang District, Daping, Heilongjiang, PRC; 55Tangshan No. 5 Hospital, Luan District, Tangshan, Hebei, PRC; 56Anshan Psychiatric Rehabilitation Hospital, Liushan District, Anshan, Liaoning, PRC; 57Weihai Mental Health Center, ETDS, Weihai, Shandong, PRC; 58Renmin Hospital of Wuhan University, Wuchang District, Wuhan, Hubei, PRC; 59Tianjin First Center Hospital, Hedong District, Tianjin, PRC; 60Hainan Anning Hospital, Hukou, Hainan, PRC

Background. The symptoms of major depression (MD) are clinically diverse. Do they form coherent factors that might clarify the underlying nature of this important psychiatric syndrome?

Method. Symptoms at lifetime worst depressive episode were assessed at structured psychiatric interview in 6008 women of Han Chinese descent, age ≥30 years with recurrent DSM-IV MD. Exploratory factor analysis (EFA) and confirmatory factor analysis (CFA) were performed in Mplus in random split-half samples.

Results. The preliminary EFA results were consistently supported by the findings from CFA. Analyses of the nine DSM-IV MD symptomatic A criteria revealed two factors loading on: (i) general depressive symptoms; and (ii) guilt/suicidal

--

* Address for correspondence: K. S. Kendler, M.D., Virginia Institute for Psychiatric and Behavioral Genetics of VCU, Box 980126, Richmond, VA 23298-0126, USA.

(E-mail: kendler@vcu.edu)

The online version of this article is published within an Open Access environment subject to the conditions of the Creative Commons Attribution license <http://creativecommons.org/licenses/by/3.0/>. 
characteristics and specific analyses.

The authors noted that this literature was inclusive evidence for any single typology of depressive symptoms. The authors noted that this literature was diverse and did not provide consensus (Kendler et al. 2010). Results were diverse and did not provide conclusive evidence for any single typology of depressive symptoms. The authors noted that this literature was difficult to interpret because of differences in sample characteristics and specific items included in the factor analyses.

This study seeks to determine the factor structure of the symptoms of MD as reported for the lifetime worst episode in a very large sample of Han Chinese women with recurrent depression. Our sample size is sufficient to utilize EFA in a random split-half of the sample. The results are then verified using CFA in the second split-half.

We performed factor analyses on three sets of items: (i) the nine DSM-IV A criteria for MD (APA, 1994); (ii) 14 items representing the disaggregated DSM-IV criteria (i.e. increased and decreased appetite and weight, insomnia or hypersomnia, and psychomotor agitation or retardation); and (iii) 27 items, consisting of the 14 above mentioned plus 13 more detailed assessments of DSM criteria and symptoms of melancholia, anxiety and Beck’s cognitive trio (Beck et al. 1980).

Method

Subjects for this report came from the China, Oxford and VCU (Virginia Commonwealth University) Experimental Research on Genetic Epidemiology (CONVERGE) study (Flint et al. 2012). The analysis was based on the final phenotypic sample of 6008 cases recruited from 57 mental health centers and psychiatric departments of general hospitals in 45 cities in 23 provinces in China.

Subjects were females aged 30–60 years who reported two or more lifetime episodes of MD meeting DSM-IV criteria (APA, 1994). Their age at onset was between 14 and 55 years and each had four Han Chinese grandparents. Cases were excluded if they had a history of mania, psychosis outside of mood episodes or mental retardation, or had abused drugs or alcohol before their first depressive episode. The samples were primarily collected for a large genome-wide association study. Twin studies have suggested that the heritability of MD is higher in females and the genetic risk factors are not identical across sexes (Kendler et al. 2001, 2006). We therefore studied women only to reduce genetic heterogeneity.
Subjects were interviewed for an average of 2h using a computerized assessment system that included assessment of psychopathology, demographic and personal characteristics, and psychosocial functioning. All interviewers, junior psychiatrists, postgraduate medical students, or senior nurses, were trained by the CONVERGE team for a minimum of 1 week. Interviews were recorded and a proportion of them edited for quality control. The study protocol was approved by the Ethical Review Board of Oxford University and the ethics committees in participating Chinese hospitals.

Measures

The diagnosis of MD was established with the Composite International Diagnostic Interview (CIDI; World Health Organization lifetime version 2.1; Chinese version), which classifies diagnoses according to DSM-IV criteria (APA, 1994). The interview was originally translated into Mandarin by a team of psychiatrists in Shanghai Mental Health Center, with the translation reviewed and modified by members of the CONVERGE team. The interview was supplemented by sections from interviews used in the Virginia Adult Twin Study of Psychiatric and Substance Use Disorders (VATSPSUD) (Kendler & Prescott, 2006), and other items of interest to the investigators.

All interview sections, with built-in skip patterns, were computerized into a bilingual system of Mandarin and English that was installed on laptops. Once an interview was completed, a backup file containing the interview data was generated and, together with the audio record of the interview, uploaded to a server in Beijing and then transferred to Oxford. As operationalized in the CIDI, the symptoms analysed were based on those reported by the subject for the self-identified worst lifetime episode of MD.

Statistical methods

The total sample was randomly divided into two halves. The first sample was used to perform an EFA and the second was used to perform a CFA for validating the EFA symptom structure. Factor analyses were performed using the Mplus program (Muthén & Muthén, 1998), with a weighted least squares means and variance adjusted (WLSMV) estimator that is designed for ordinal data (Flora & Curran, 2004). EFA was performed using a geomin oblique rotation. CFA model fit was evaluated using the Tucker–Lewis index (TLI; Tucker & Lewis, 1973), the comparative fit index (CFI; Bentler, 1990) and the root mean square error of approximation (RMSEA; Steiger, 1990). For the TLI and CFI, values between 0.90 and 0.95 are considered acceptable, and ≥0.95 as good. For the RMSEA, good models have values ≤0.05. To confirm that the split-sample procedure produced two random subsamples, we formally tested in Mplus for measurement invariance in our CFA data on the nine DSM-IV criteria. When we constrained factor loadings to be identical in the two subsamples, the robust chi-square difference, as expected, was not significant.

Results

EFA and CFA on nine MD DSM-IV A criteria

Endorsement frequencies for the nine MD DSM-IV A criteria for MD were uniformly high and similar in our two subsamples (Table 1). An EFA in sample 1 produced two eigenvalues exceeding unity. The first factor displayed higher loadings primarily on the somatic symptoms of depression and less prominently on mood symptoms, and reflected a ‘general depressive symptom’ factor. The second factor had highest loadings on suicidal ideation and worthlessness/guilt, and reflected a ‘cognitive symptoms’ factor (Table 1). The inter-factor correlation coefficient was +0.53. The best-fit CFA solution, two factors with all variable loadings >0.30 from the EFA, had an excellent fit (RMSEA=0.015, CFI=0.98, TLI=0.97), which was substantially better than those obtained with one factor (RMSEA=0.031, CFI=0.89, TLI=0.86). The inter-factor correlation was +0.78. The loadings for the mood criteria A1 was stronger on the first factor and much weaker on the second factor in the CFA compared with the EFA. This factor, which still represented ‘general depressive symptoms’, had high loadings on seven of the nine A criteria. Factor 2, the ‘cognitive symptoms’ factor, had high loadings on the two criteria of worthlessness/guilt and suicidal ideation.

EFA and CFA on the 14 disaggregated DSM-IV A criteria

The frequencies of item endorsement in the 14 disaggregated DSM-IV A criteria were similar in the two subsamples and more variable than with the nine criteria (Table 2). Atypical vegetative symptoms were endorsed by 7–13% of patients. The scree plot for the EFA analyses showed four factors with eigenvalues exceeding 1 and a clear ‘elbow’ at three factors. Furthermore, a solution with three factors was considerably more interpretable than one with four and resulted in factors reflecting ‘weight/appetite symptoms’, ‘general depressive symptoms’ and ‘sleep disturbance’ (Table 2) with modest inter-factor correlations (Table 3). In our CFA analyses (which produced the most interpretable results when including all EFA loadings ≥0.30), the three-factor solution had better-fit...
indices (RMSEA=0.032, CFI=0.94, TLI=0.93) than the four-factor solution (RMSEA=0.035, CFI=0.93, TLI=0.91), validating our interpretation of the EFA. Loadings from the CFA were comparable with those found in the EFA, identifying three broadly similar factors. Inter-factor correlations were modest except between factors 1 and 3, which equaled −0.44 (Table 3).

### Table 1. Item percentage endorsement frequency on sample 1 and sample 2 for the nine DSM-IV criteria for major depression and loadings estimated from an EFA on sample 1 and a CFA on sample 2

<table>
<thead>
<tr>
<th>Items</th>
<th>EFA on sample 1</th>
<th>CFA on sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Endorsement, %</td>
<td>Factor 1</td>
</tr>
<tr>
<td>Depressed mood</td>
<td>99.57</td>
<td>0.363</td>
</tr>
<tr>
<td>Anhedonia</td>
<td>98.87</td>
<td>0.266</td>
</tr>
<tr>
<td>Significant weight or appetite disturbance</td>
<td>91.08</td>
<td>0.511</td>
</tr>
<tr>
<td>Sleep disturbance</td>
<td>95.67</td>
<td>0.539</td>
</tr>
<tr>
<td>Psychomotor agitation or retardation</td>
<td>91.61</td>
<td>0.647</td>
</tr>
<tr>
<td>Loss of energy or fatigue</td>
<td>93.53</td>
<td>0.719</td>
</tr>
<tr>
<td>Feeling of worthlessness</td>
<td>90.67</td>
<td>0.106</td>
</tr>
<tr>
<td>Diminished ability to think or concentrate, or indecisiveness</td>
<td>97.72</td>
<td>0.661</td>
</tr>
<tr>
<td>Recurrent thoughts of death</td>
<td>76.32</td>
<td>−0.011</td>
</tr>
</tbody>
</table>

DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, fourth edition; EFA, exploratory factor analysis; CFA, confirmatory factor analysis.

* Two-factor solution with loading cut-off at 0.3.

### Table 2. Item percentage endorsement frequency on sample 1 and sample 2 for the 14 disaggregated DSM-IV criteria for major depression and loadings estimated from an EFA on sample 1 and a CFA on sample 2

<table>
<thead>
<tr>
<th>Items</th>
<th>EFA on sample 1</th>
<th>CFA on sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Endorsement, %</td>
<td>Factor 1</td>
</tr>
<tr>
<td>Depressed mood</td>
<td>99.57</td>
<td>0.093</td>
</tr>
<tr>
<td>Anhedonia</td>
<td>98.87</td>
<td>−0.082</td>
</tr>
<tr>
<td>Loss of appetite</td>
<td>84.23</td>
<td>−0.722</td>
</tr>
<tr>
<td>Loss of weight</td>
<td>60.15</td>
<td>−0.621</td>
</tr>
<tr>
<td>Increase of appetite</td>
<td>9.43</td>
<td>0.738</td>
</tr>
<tr>
<td>Increase of weight</td>
<td>7.11</td>
<td>0.842</td>
</tr>
<tr>
<td>Insomnia</td>
<td>92.99</td>
<td>−0.003</td>
</tr>
<tr>
<td>Hypersomnia</td>
<td>13.12</td>
<td>0.016</td>
</tr>
<tr>
<td>Psychomotor retardation</td>
<td>76.69</td>
<td>−0.105</td>
</tr>
<tr>
<td>Psychomotor agitation</td>
<td>72.19</td>
<td>−0.008</td>
</tr>
<tr>
<td>Loss of energy or fatigue</td>
<td>93.53</td>
<td>−0.100</td>
</tr>
<tr>
<td>Feeling of worthlessness</td>
<td>90.67</td>
<td>0.048</td>
</tr>
<tr>
<td>Diminished ability to think or concentrate, or indecisiveness</td>
<td>97.72</td>
<td>0.015</td>
</tr>
<tr>
<td>Recurrent thoughts of death</td>
<td>76.32</td>
<td>−0.074</td>
</tr>
</tbody>
</table>

DSM-IV, Diagnostic and Statistical Manual of Mental Disorders, fourth edition; EFA, exploratory factor analysis; CFA, confirmatory factor analysis.

* Items selected to load on CFA factors are based on EFA loading with a cut-off of 0.3.
Endorsement rates were variable for the 27 depression-related symptoms individually assessed during the worst lifetime episode and similar in our two subsamples (Table 4). More than 80% of our sample endorsed both key melancholic symptoms (e.g. lack of mood reactivity and distinct mood quality) and items assessing Beck’s cognitive triad (e.g. hopelessness and worthlessness) indicating the clinical severity of the reported depressive episodes.

The factor structure of these 27 items was initially examined by EFA (Table 3), resulting in eight eigenvalues exceeding 1 without a clear ‘elbow’ in the scree plot. With seven- and eight-factor solutions, estimated residual variances of some items became negative, indicating over-extraction. The five-factor EFA model was clinically more sensible than the six-factor solution. Similar to the results from the 14-item analysis, the first factor reflected ‘weight/appetite symptoms’ (Table 4). The second factor was also similar to that seen with the 14-item solution with a more prominent loading on psychomotor retardation. We called this the ‘general retarded depressive symptom’ factor. The third factor had highest loadings on the ‘atypical vegetative symptoms’ of increased appetite and weight, and hypersomnia. The fourth factor had prominent loadings on suicidal/hopeless symptoms including agitation, nervousness, guilt, irritability and crying. Inter-factor correlations were generally modest (Table 3), with the highest observed between the general depressive and suicidal/hopeless factors (+0.38).

In the CFA, when the number of factors equalled six or greater, standard errors of model parameter estimates as well as the factor scores could not be computed, which is also an indication of factor over-extraction. Both four-factor and five-factor solutions (with items of EFA loading cut-off of 0.2) described the data well, with the five-factor solution resulting in slightly superior fits (RMSEA=0.030, CFI=0.953, TLI=0.946) compared with the four-factor solution (RMSEA=0.030, CFI=0.950, TLI=0.943), again congruent with our clinical interpretation. The loadings closely resembled those found in the EFA. The general retarded depressive symptom factor had notable loadings (with absolute value ≥0.2) on by far the most items (n=15) with higher loadings for typical severe retarded depressive symptoms including several of the melancholic symptoms. The agitated depressive and suicidal/hopeless factors were next, with prominent loadings on six and five symptoms, respectively. The atypical vegetative and weight/appetite symptom factors were the smallest, with prominent loadings on four and two symptoms, respectively.

**Discussion**

The goal of this study was to clarify the structure of the symptoms of MD experienced during the lifetime worst depressive episodes reported by Han Chinese women with recurrent DSM-IV MD (APA, 1994) ascertained in clinical settings throughout China. We conducted an EFA in a random split-half of the sample...
Table 4. Item percentage endorsement frequency on sample 1 and sample 2 for the 27 depressive symptoms and loadings estimated from an EFA on sample 1 and a CFA on sample 2.

<table>
<thead>
<tr>
<th>Items</th>
<th>EFA on sample 1</th>
<th>CFA on sample 2</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Endorsement, %</td>
<td>Factor 1</td>
</tr>
<tr>
<td>Depressed mood</td>
<td>99.57</td>
<td>-0.055</td>
</tr>
<tr>
<td>Anhedonia</td>
<td>98.87</td>
<td>0.025</td>
</tr>
<tr>
<td>Loss of appetite</td>
<td>84.23</td>
<td>0.710</td>
</tr>
<tr>
<td>Loss of weight</td>
<td>60.15</td>
<td>0.647</td>
</tr>
<tr>
<td>Increase of appetite</td>
<td>9.43</td>
<td>-0.653</td>
</tr>
<tr>
<td>Increase of weight</td>
<td>7.11</td>
<td>-0.742</td>
</tr>
<tr>
<td>Insomnia</td>
<td>92.99</td>
<td>0.061</td>
</tr>
<tr>
<td>Hypersomnia</td>
<td>13.12</td>
<td>-0.091</td>
</tr>
<tr>
<td>Psychomotor retardation</td>
<td>76.69</td>
<td>0.038</td>
</tr>
<tr>
<td>Psychomotor agitation</td>
<td>72.19</td>
<td>-0.033</td>
</tr>
<tr>
<td>Loss of energy or fatigue</td>
<td>93.53</td>
<td>0.116</td>
</tr>
<tr>
<td>Worthlessness</td>
<td>82.39</td>
<td>-0.016</td>
</tr>
<tr>
<td>Guilty or remorse</td>
<td>72.42</td>
<td>0.018</td>
</tr>
<tr>
<td>Trouble concentrating</td>
<td>92.15</td>
<td>-0.056</td>
</tr>
<tr>
<td>Slower/mixed-up thoughts</td>
<td>89.30</td>
<td>-0.180</td>
</tr>
<tr>
<td>Trouble making decision</td>
<td>83.36</td>
<td>-0.099</td>
</tr>
<tr>
<td>Death thoughts</td>
<td>71.22</td>
<td>-0.010</td>
</tr>
<tr>
<td>Suicidal thoughts</td>
<td>61.72</td>
<td>-0.067</td>
</tr>
<tr>
<td>Loss of ability to enjoy good things</td>
<td>87.05</td>
<td>0.042</td>
</tr>
<tr>
<td>Depression different from grief or loss</td>
<td>93.54</td>
<td>0.103</td>
</tr>
<tr>
<td>Worse mood in the morning</td>
<td>63.70</td>
<td>0.019</td>
</tr>
<tr>
<td>Loss of sexual drive</td>
<td>89.21</td>
<td>0.020</td>
</tr>
<tr>
<td>Feel irritable or angry most of the time</td>
<td>74.59</td>
<td>-0.065</td>
</tr>
<tr>
<td>Hopeless</td>
<td>80.80</td>
<td>0.068</td>
</tr>
<tr>
<td>Cry a lot</td>
<td>67.55</td>
<td>0.140</td>
</tr>
<tr>
<td>Helpless</td>
<td>89.65</td>
<td>0.113</td>
</tr>
<tr>
<td>Feel nervous, jittery or anxious</td>
<td>89.58</td>
<td>-0.032</td>
</tr>
</tbody>
</table>

EFA, Exploratory factor analysis; CFA, Confirmatory factor analysis.

* Items selected to load on CFA factors are based on EFA loading with a cut-off of 0.2.
and then verified those results by CFA in the second random half. We conducted these analyses with:
(i) the nine DSM-IV A criteria for MD; (ii) 14 items representing the disaggregated DSM-IV A criteria;
and (iii) all 27 symptoms independently assessed in our interview for the worst lifetime depressive episode.
Our preliminary EFA results on one half of our sample were consistently verified by the findings from CFA
in the second random half.

Our study produced three major findings. First, we detected two factors underlying the nine DSM
criteria for MD. The two ‘cognitive’ DSM criteria—worthlessness/guilt and suicidal ideation formed their
own factor. Our results are convergent with prior analyses from the VATSPSUD (Lux & Kendler, 2010) that
detected evidence for covert heterogeneity within the DSM-IV A criteria for MD, particularly between
somatic and cognitive criteria (Lux & Kendler, 2010), as well as results from the Beck depression rating
scale that cognitive symptoms form a factor independent of the somatic and affective features of depression
(Beck et al. 1961). These findings suggest that in clinically depressed individuals, the DSM-IV A criteria
for MD—which have changed very little since the days of the Research Diagnostic Criteria (Spitzer et al.
1975) and DSM-III (APA, 1980)—in fact reflect two inter-correlated but meaningfully distinct rather than
one underlying psychopathological dimension.

It is particularly useful to compare our findings with
those reported by Van Loo et al. (2012) who reviewed
the results for nine studies of EFA and principal com-
ponent analysis of DSM MD criteria. Of note, the
sample size in our study is substantially greater than
that of all nine of these prior studies. Contrary to the
report of Aggen et al. (2005), all the studies reviewed
detected at least two factors. Most commonly, as seen
in our CFA, the core items of sad mood and loss of
interest loaded on the first factor. However, the other
features of the results of the individual studies were
quite varied, further illustrating how dependent the
results of factor analysis can be on the nature of the
items examined and the sample studied.

Second, consistent with several prior multivariate
analyses of MD criteria (Kendler et al. 1996; Sullivan
et al. 1998; Matza et al. 2003) and depressive symptom
scales (Bech et al. 2011), when atypical vegetative
symptoms are well represented among the items, they form
a distinct dimension of depressive symptomatology
(Stewart et al. 1993). However, contrary to most prior
studies of clinical criteria, perhaps because of our
statistical power, we found that items reflecting sleep
difficulty were moderately independent from the other
vegetative symptoms reflecting changes in weight/
appetite and formed their own factor. However, our
results are consistent with a meta-analysis of the
Hamilton Rating Scale for Depression that demonstra-
ted separate factors in clinically depressed subjects
for sleep problems and weight loss (Shafer, 2006).
Sleep disturbances in severe MD may form an important
psychopathological dimension at least partially inde-
pendent of other neurovegetative changes (Kupfer,
1995).

Third, when we added to the DSM-IV disaggregated
A criteria a range of other common depressive symp-
toms, a richer and potentially more informative picture
emerged of the clinical complexity of severe depressive
illness. We found replicated evidence for five de-
pressive symptom factors. The largest factor contained
prominent loadings on traditional mood changes, psy-
chomotor retardation and several melancholic criteria,
with modest loadings on certain cognitive symptoms.
Items in this general retarded depressive symptom
factor reflected six of the nine DSM IV A criteria: sad
mood, anhedonia, psychomotor changes, worthless-
ness, fatigue, and concentration problems.

The next largest factor reflected agitated and anxious
depressive symptoms, with additional high loadings
on guilt, crying and feelings of helplessness. These
symptoms resemble those describing the agitated
depression (Koukopoulos & Koukopoulos, 1999)
that was recognized as a subtype in the Research Dia-
nostic Criteria (Spitzer et al. 1975). These findings pro-
vide further evidence that anxiety symptoms can be
a prominent part of the presentation of some forms
of severe MD and may form a potentially important
independent symptom factor as has been proposed
for DSM-5 (First, 2011).

Perhaps the most interesting factor to emerge had
prominent loadings on suicidal symptoms and Beck’s
triad of hopelessness, worthlessness and (somewhat
more weakly) helplessness (Beck & Alford, 2008).
Consistent with a large body of influential work
(Beck et al. 1980; Beck & Alford, 2008), our results
suggest that cognitive symptoms reflecting distorted
views of the self, the world and the future form an
important and largely independent dimension of depressive symptomatology which in this ascertained
population is very closely related to suicidal ideation.
These results also suggest that criterion A7 for MD
does not reflect a unitary psychopathological con-
struct, as its two major subcomponents—worthlessness
and guilt—loaded on different depressive symptom
factors.

Finally, similar to the analyses of the 14 items, the
factor analyses of our 27 items identified two factors
reflecting typical and atypical vegetative symptoms
with weight/appetite and sleep items loading on separ-
ate factors.

Both our EFA and CFA utilized oblique rotations
so that the depressive symptoms factors could
inter-correlate. This approach produces a picture that, while more complex, is probably more realistic. For technical reasons (because CFA sets to zero low item loadings which partly capture inter-factor associations in an EFA), inter-factor correlations are typically higher in CFA than EFA. Because psychiatric symptoms are likely to be ‘messy’ and truly reflect more than one underlying dimension, the inter-factor correlations estimated with the more complex EFA factor loading patterns are probably more accurate than those produced by the stronger simple structure imposed in the CFAs. We would therefore emphasize only three of these correlations, all positive, as being especially noteworthy: between the general depressive and cognitive symptom factors in our nine-item analyses and, in our 27-item analyses between our general retarded depressive symptom factor and both the suicidal/hopeless and agitated depressive symptom factors.

Our results differ from two prior factor analyses of the nine DSM A criteria for MD in community samples (Muthén, 1989; Aggen et al. 2005), both of which detected only a single factor with substantially higher loadings on all criteria. Perhaps our ability to detect a second cognitive factor is related to our larger sample size and the greater clinical severity of our subjects. We are not aware of prior studies using similar methods and symptom content for our 14- and 27-item analyses. However, it is helpful to contextualize our findings by reviewing selectively earlier and more recent attempts to derive symptomatic factors from clinically depressed individuals. A classic factor analytic study of Grinker et al. (1961) of a wide variety of symptoms in 96 hospitalized depressed patients produced two factors similar to those we detected dominated by loadings on: (i) tension, anxiety and feeling jittery; and (ii) feeling hopeless, helpless, unworthy, and failure (Grinker et al. 1961). The initial description of the Hamilton Rating Scale for Depression reported on 49 depressed patients four symptom factors, three of which resemble those we recovered including factors dominated by high loadings on: (i) suicide and guilt; (ii) agitation and anxiety; and (iii) insomnia and weight loss (Hamilton, 1960). Friedman, in another early study of 170 cases of psychotic depression using Grinker’s 60-item depression rating scale, reported four symptom-dominated factors, three of which resembled factors that we isolated with high loadings on: (i) guilt and low self-esteem; (ii) appetite and sleep disturbance; and (iii) irritability, physical complaints and agitation (Friedman et al. 1963). An early analysis of the Beck Depression Scale on 254 hospitalized depressed patients yielded three interpretable factors of which two resembled those we had found with prominent loadings on: (i) guilt, self-accusation, sense of failure, self-punitive wishes and self-hate; and (ii) weight loss, loss of appetite, and sleep disturbance (Weckowicz et al. 1967).

Turning to more recent studies, in a meta-analysis of four popular self-report depression questionnaires, Shafer (2006) noted the consistency with which neurovegetative symptoms of depression loaded on an independent factor from other clinical symptoms. Bech et al. (2011), using principal component analysis applied to the 17-item Hamilton Rating Scale for Depression (HAMD17) scores from 4041 patients from the STAR*D (Sequenced Treatment Alternatives to Relieve Depression) study, found two factors, one reflecting general depressive severity and the second vegetative symptoms of weight, appetite and sleep disturbance. They also analysed the Inventory of Depressive Symptomatology (IDS-C) in the same sample also reporting two factors. The first factor again included a broad array of depressive symptoms, while the second factor had notable loadings on panic, arousal, agitation, but also sleep difficulties (Parker, 2007). Finally, in 1049 primary care patients with MD, Romero et al. (2008) found four factors in the Zung Depression Scale, three of which were quite similar to factors we extracted with notable loadings on: (i) emptiness, hopelessness and suicidal rumination; (ii) agitation, irritability and crying spells; and (iii) decreased appetite and weight loss.

Our subjects were Han Chinese women with recurrent and severe MD recruited through clinical settings. We cannot be sure of the degree to which these results would extrapolate to men, to other ethnic groups, or to the milder forms of illness or symptom dimensions studied in community samples. However, the World Health Organization examined patients seeking care for depression in five international sites including East Asia and concluded that the clinical similarities of individuals far outweighed the modest cross-cultural differences (Sartorius et al. 1980). The factor structure of the Beck Depression Scale in Japan is very similar to that seen in Western populations (Kojima et al. 2002). Prior studies in this sample have shown that MD is associated, in a similar manner, with a range of risk factors previously demonstrated in Western samples including childhood sexual abuse (Cong et al. 2011), neuroticism (Xia et al. 2011), stressful life events (Tao et al. 2011) and low parental warmth (Gao et al. 2012), and has similar patterns of co-morbidity with anxiety disorders (Li et al. 2012) and dysthymia (Sang et al. 2011).

Epidemiological studies have found that rates of MD are lower in East Asia than in most other countries (Parker et al. 2001). There has been supporting evidence that the Chinese tend to deny depression or express it somatically (Kleinman, 1982; Parker et al. 2001).
The leading theory is that this is a result of cultural stoicism and high levels of stigmatization. A recent epidemiological study in Taiwan supports this hypothesis. Individuals reporting MD were much more impaired than subjects reporting MD in parallel US studies and were only one-third as likely to seek professional help (Liao et al. 2012). This might also explain the clinical severity of our sample that was ascertained entirely through psychiatric treatment centers.

While it is possible that the patterns of depressive illness described in our sample are unique to this subgroup, this seems unlikely given prior cross-cultural research and the many parallels between our findings and previous empirical studies of MD in Western populations.

To our knowledge, this is the first study to examine the underlying symptom pattern of depression in the Han Chinese population using a carefully screened clinical sample of this size. The fact that we have identified common factors from Western samples provides evidence that there is more similarity than dissimilarities in depression in China and in Western countries. Future studies are needed to compare samples recruited using the same criteria from diverse ethnic groups to address more definitively the role played by culture and ethnicity in shaping the symptom dimensions of depression.

**Strengths and limitations**

These results should be considered in the light of our potential methodological strengths and limitations. Four strengths are noteworthy. First, our phenotypic assessment was standardized and detailed with careful interview training and several quality-control features. Second, our sample was much larger than any previous similar effort and allowed us, with good power, to utilize EFA and CFA methods in split-half samples. Third, the goal of our sample collection strategy was to minimize heterogeneity that could add noise to our analysis. So, this sample was of a single ethnicity who suffered from severe recurrent depressive illness. Furthermore, drug and alcohol abuse were vanishingly rare in our sample of Han Chinese women, reducing a further major source of confound; also only 5% of them ever smoked. Fourth, the sample had a minimum age of 30 years and a mean age of 44.4 (s.d. = 8.9) years, reducing substantially the proportion that might eventually develop bipolar illness.

One major methodological limitation deserves comment. The results of factor analysis are critically dependent on the items analysed. We utilized a standardized interview for the DSM-IV criteria both in aggregated and disaggregated form. However, the other items added to form our final 27-item analysis reflected our own research interests. Furthermore, we could not include all the DSM-IV melancholic symptoms because some of them just assess the same symptoms as one of the A criteria at a more stringent level (e.g., melancholic criterion B4 is just a severe version of criterion A5). Including both items would cause estimation problems because the melancholic criterion would never be scored present in the absence of the relevant A criterion.

**Acknowledgements**

This work was funded by the Wellcome Trust. All authors are part of the CONVERGE consortium (China, Oxford and VCU Experimental Research on Genetic Epidemiology) and gratefully acknowledge the support of all partners in hospitals across China.

**Declaration of Interest**

None.

**References**


